



Sequence Listing to File.TXT
SEQUENCE LISTING

<110> YOSHINAGA, Takashi
ARAI, Toru

<120> hERG channel-expressing cell

<130> 09857/0204372-US0

<140> 10/595,858
<141> 2006-05-16

<150> PCT/JP2004/017441
<151> 2004-11-17

<150> JP 2003-387255
<151> 2003-11-17

<160> 13

<170> PatentIn version 3.3

<210> 1
<211> 4070
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (184)..(3660)

<400> 1
acgcggcctg ctcaggcctc cagcgccgg tcggagggga ggcgggaggc gagcgaggac 60
ccgcgcccgc agtccagtct gtgcgcgcc gtcgtcgctt ggcgcgggtgc gggaccagcg 120
ccggccaccc gaagcctagt gcgtcgccgg gtgggtgggc ccgcggcgcc ccatgggctc 180
agg atg ccg gtg cgg agg ggc cac gtc gcg ccg cag aac acc acc ttc ctg 228
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu
1 5 10 15
gac acc atc atc cgc aag ttt gag ggc cag agc cgt aag ttc atc atc 276
Asp Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Ile Ile
20 25 30
gcc aac gct cgg gtg gag aac tgc gcc gtc atc tac tgc aac gac ggc 324
Ala Asn Ala Arg Val Glu Asn Cys Ala Val Ile Tyr Cys Asn Asp Gly
35 40 45
ttc tgc gag ctg tgc ggc tac tcg cgg gcc gag gtg atg cag cga ccc 372
Phe Cys Glu Leu Cys Gly Tyr Ser Arg Ala Glu Val Met Gln Arg Pro
50 55 60
tgc acc tgc gac ttc ctg cac ggg cgc acg cag cgc cgc gct gcc 420
Cys Thr Cys Asp Phe Leu His Gly Pro Arg Thr Gln Arg Arg Ala Ala
65 70 75
gcg cag atc gcg cag gca ctg ctg ggc gcc gag gag cgc aaa gtg gaa 468
Ala Gln Ile Ala Gln Ala Leu Leu Gly Ala Glu Glu Arg Lys Val Glu
80 85 90 95

Sequence Listing to File.TXT

| | |
|--|------|
| atc gcc ttc tac cg ^g aaa gat gg ^g agc tgc ttc cta tgt ctg gtg gat Ile Ala Phe Tyr Arg Lys Asp Gly Ser Cys Phe Leu Cys Leu Val Asp | 516 |
| 100 105 110 | |
| gtg gtg ccc gtg aag aac gag gat gg ^g gct gtc atc atg ttc atc ctc Val Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu | 564 |
| 115 120 125 | |
| aat ttc gag gtg gtg atg gag aag gac atg gtg gg ^g tcc ccg gct cat Asn Phe Glu Val Val Met Glu Lys Asp Met Val Gly Ser Pro Ala His | 612 |
| 130 135 140 | |
| gac acc aac cac cg ^g gg ^c ccc ccc acc agc tgg ctg gcc cca gg ^c cgc Asp Thr Asn His Arg Gly Pro Pro Thr Ser Trp Leu Ala Pro Gly Arg | 660 |
| 145 150 155 | |
| gcc aag acc ttc cg ^c ctg aag ctg ccc gc ^g ctg ctg gc ^g ctg acg gcc Ala Lys Thr Phe Arg Leu Lys Leu Pro Ala Leu Leu Ala Leu Thr Ala | 708 |
| 160 165 170 175 | |
| cg ^g gag tcg tcg gtg cg ^g tcg gg ^c gg ^c gc ^g gg ^c gc ^g gg ^c gc ^g Arg Glu Ser Ser Val Arg Ser Gly Gly Ala Gly Gly Ala Gly Ala Pro | 756 |
| 180 185 190 | |
| gg ^g gcc gtg gtg gac gtg gac ctg acg ccc gc ^g gca ccc agc agc Gly Ala Val Val Val Asp Val Asp Leu Thr Pro Ala Ala Pro Ser Ser | 804 |
| 195 200 205 | |
| gag tcg ctg gcc ctg gac gaa gtg aca gcc atg gac aac cac gtg gca Glu Ser Leu Ala Leu Asp Glu Val Thr Ala Met Asp Asn His Val Ala | 852 |
| 210 215 220 | |
| gg ^g ctc gg ^c ccc gc ^g gag gag cg ^g cgt gc ^g ctg gtg ggt ccc gg ^c tct Gly Leu Gly Pro Ala Glu Glu Arg Arg Ala Leu Val Gly Pro Gly Ser | 900 |
| 225 230 235 | |
| ccg ccc cgc agc gc ^g ccc gg ^c cag ctc cca tcg ccc cgg gc ^g cac agc Pro Pro Arg Ser Ala Pro Gly Gln Leu Pro Ser Pro Arg Ala His Ser | 948 |
| 240 245 250 255 | |
| ctc aac ccc gac gg ^c tcg gg ^c tcc agc tgc agc ctg gcc cgg acg cgc Leu Asn Pro Asp Ala Ser Gly Ser Ser Cys Ser Leu Ala Arg Thr Arg | 996 |
| 260 265 270 | |
| tcc cga gaa agc tgc gg ^c agc gtg cgc cgc gcc tcg tcg gcc gac gac Ser Arg Glu Ser Cys Ala Ser Val Arg Arg Ala Ser Ser Ala Asp Asp | 1044 |
| 275 280 285 | |
| atc gag gg ^c atg cg ^c gg ^g gtg ctg ccc ccg cca ccg gc ^g cac gg ^c Ile Glu Ala Met Arg Ala Gly Val Leu Pro Pro Pro Arg His Ala | 1092 |
| 290 295 300 | |
| agc acc gg ^c gg ^c atg cac cca ctg cgc agc gg ^c ttg ctc aac tcc acc Ser Thr Gly Ala Met His Pro Leu Arg Ser Gly Leu Leu Asn Ser Thr | 1140 |
| 305 310 315 | |
| tcg gac tcc gac ctc gtg cgc tac cgc acc att agc aag att ccc caa Ser Asp Ser Asp Leu Val Arg Tyr Arg Thr Ile Ser Lys Ile Pro Gln | 1188 |
| 320 325 330 335 | |
| atc acc ctc aac ttt gtg gac ctc aag gg ^c gac ccc ttc ttg gct tcg Ile Thr Leu Asn Phe Val Asp Leu Lys Gly Asp Pro Phe Leu Ala Ser | 1236 |
| 340 345 350 | |

Sequence Listing to File.TXT

| | |
|---|------|
| ccc acc agt gac cgt gag atc ata gca cct aag ata aag gag cga acc Pro Thr Ser Asp Arg Glu Ile Ile Ala Pro Lys Ile Lys Glu Arg Thr 355 360 365 | 1284 |
| cac aat gtc act gag aag gtc acc cag gtc ctg tcc ctg ggc gcc gac His Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala Asp 370 375 380 | 1332 |
| gtg ctg cct gag tac aag ctg cag gca ccg cgc atc cac cgc tgg acc Val Leu Pro Glu Tyr Lys Leu Gln Ala Pro Arg Ile His Arg Trp Thr 385 390 395 | 1380 |
| atc ctg cat tac agc ccc ttc aag gcc gtg tgg gac tgg ctc atc ctg Ile Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile Leu 400 405 410 415 | 1428 |
| ctg ctg gtc atc tac acg gct gtc ttc aca ccc tac tcg gct gcc ttc Leu Leu Val Ile Tyr Thr Ala Val Phe Thr Pro Tyr Ser Ala Ala Phe 420 425 430 | 1476 |
| ctg ctg aag gag acg gaa gaa ggc ccg cct gct acc gag tgt ggc tac Leu Leu Lys Glu Thr Glu Glu Gly Pro Pro Ala Thr Glu Cys Gly Tyr 435 440 445 | 1524 |
| gcc tgc cag ccg ctg gct gtg gtg gac ctc atc gtg gac atc atg ttc Ala Cys Gln Pro Leu Ala Val Val Asp Leu Ile Val Asp Ile Met Phe 450 455 460 | 1572 |
| att gtg gac atc ctc atc aac ttc cgc acc acc tac gtc aat gcc aac Ile Val Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Ala Asn 465 470 475 | 1620 |
| gag gag gtg gtc agc cac ccc ggc ccg atc gcc gtc cac tac ttc aag Glu Glu Val Val Ser His Pro Gly Arg Ile Ala Val His Tyr Phe Lys 480 485 490 495 | 1668 |
| ggc tgg ttc ctc atc gac atg gtg gcc gcc atc ccc ttc gac ctg ctc Gly Trp Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu 500 505 510 | 1716 |
| atc ttc ggc tct ggc tct gag gag ctg atc ggg ctg ctg aag act gcg Ile Phe Gly Ser Gly Ser Glu Glu Leu Ile Gly Leu Leu Lys Thr Ala 515 520 525 | 1764 |
| cgg ctg ctg cgg ctg gtg cgc gtg gcg ccg aag ctg gat cgc tac tca Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser 530 535 540 | 1812 |
| gag tac ggc gcg gcc gtg ctg ttc ttg ctc atg tgc acc ttt gcg ctc Glu Tyr Gly Ala Ala Val Leu Phe Leu Leu Met Cys Thr Phe Ala Leu 545 550 555 | 1860 |
| atc gcg cac tgg cta gcc tgc atc tgg tac gcc atc ggc aac atg gag Ile Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Met Glu 560 565 570 575 | 1908 |
| cag cca cac atg gac tca cgc atc ggc tgg ctg cac aac ctg ggc gac Gln Pro His Met Asp Ser Arg Ile Gly Trp Leu His Asn Leu Gly Asp 580 585 590 | 1956 |
| cag ata ggc aaa ccc tac aac agc agc ggc ctg ggc ggc ccc tcc atc Gln Ile Gly Lys Pro Tyr Asn Ser Ser Gly Leu Gly Gly Pro Ser Ile | 2004 |

Sequence Listing to File.TXT

595

600

605

| | | | |
|---|-----|-----|------|
| aag gac aag tat gtg acg gcg ctc tac ttc acc ttc agc agc ctc acc Lys Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr 610 | 615 | 620 | 2052 |
| agt gtg ggc ttc ggc aac gtc tct ccc aac acc aac tca gag aag atc Ser Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Ile 625 | 630 | 635 | 2100 |
| ttc tcc atc tgc gtc atg ctc att ggc tcc ctc atg tat gct agc atc Phe Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile 640 | 645 | 650 | 2148 |
| ttc ggc aac gtg tcg gcc atc atc cag cg ^g ctg tac tcg ggc aca gcc Phe Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala 660 | 665 | 670 | 2196 |
| cgc tac cac aca cag atg ctg cg ^g gtg cg ^g gag ttc atc cgc ttc cac Arg Tyr His Thr Gln Met Leu Arg Val Arg Glu Phe Ile Arg Phe His 675 | 680 | 685 | 2244 |
| cag atc ccc aat ccc ctg cgc cag cgc ctc gag gag tac ttc cag cac Gln Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His 690 | 695 | 700 | 2292 |
| gcc tgg tcc tac acc aac ggc atc gac atg aac gcg gtg ctg aag ggc Ala Trp Ser Tyr Thr Asn Gly Ile Asp Met Asn Ala Val Leu Lys Gly 705 | 710 | 715 | 2340 |
| ttc cct gag tgc ctg cag gct gac atc tgc ctg cac ctg aac cgc tca Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn Arg Ser 720 | 725 | 730 | 2388 |
| ctg ctg cag cac tgc aaa ccc ttc cga ggg gcc acc aag ggc tgc ctt Leu Leu Gln His Cys Lys Pro Phe Arg Gly Ala Thr Lys Gly Cys Leu 740 | 745 | 750 | 2436 |
| cgg gcc ctg gcc atg aag ttc aag acc aca cat gca ccg cca ggg gac Arg Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Pro Pro Gly Asp 755 | 760 | 765 | 2484 |
| aca ctg gtg cat gct ggg gac ctg ctc acc gcc ctg tac ttc atc tcc Thr Leu Val His Ala Gly Asp Leu Leu Thr Ala Leu Tyr Phe Ile Ser 770 | 775 | 780 | 2532 |
| cgg ggc tcc atc gag atc ctg cg ^g ggc gac gtc gtc gtg gcc atc ctg Arg Gly Ser Ile Glu Ile Leu Arg Gly Asp Val Val Val Ala Ile Leu 785 | 790 | 795 | 2580 |
| ggg aag aat gac atc ttt ggg gag cct ctg aac ctg tat gca agg cct Gly Lys Asn Asp Ile Phe Gly Glu Pro Leu Asn Leu Tyr Ala Arg Pro 800 | 805 | 810 | 2628 |
| ggc aag tcg aac ggg gat gtg cg ^g gcc ctc acc tac tgt gac cta cac Gly Lys Ser Asn Gly Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His 820 | 825 | 830 | 2676 |
| aag atc cat cgg gac gac ctg ctg gag gtg ctg gac atg tac cct gag Lys Ile His Arg Asp Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Glu 835 | 840 | 845 | 2724 |
| ttc tcc gac cac ttc tgg tcc agc ctg gag atc acc ttc aac ctg cga | | | 2772 |

Sequence Listing to File.TXT

| | | | |
|---|---------------------|-----------------|------|
| Phe Ser Asp His Phe Trp Ser | Ser Leu Glu Ile Thr | Phe Asn Leu Arg | 860 |
| 850 | 855 | | |
| gat acc aac atg atc ccg ggc tcc ccc ggc agt acg gag tta gag ggt | | | 2820 |
| Asp Thr Asn Met Ile Pro Gly Ser Pro Gly Ser Thr Glu Leu Glu Gly | | | |
| 865 | 870 | 875 | |
| ggc ttc agt cgg caa cgc aag cgc aag ttg tcc ttc cgc agg cgc acg | | | 2868 |
| Gly Phe Ser Arg Gln Arg Lys Arg Lys Leu Ser Phe Arg Arg Arg Thr | | | |
| 880 | 885 | 890 | 895 |
| gac aag gac acg gag cag cca ggg gag gtg tcg gcc ttg ggg ccg ggc | | | 2916 |
| Asp Lys Asp Thr Glu Gln Pro Gly Glu Val Ser Ala Leu Gly Pro Gly | | | |
| 900 | 905 | | 910 |
| cg ^g gc ^g gg ^g gc ^a gg ^g cc ^g ag ^t ag ^c cc ^g gg ^c cc ^g gg ^g gg ^g cc ^g tt ^g | | | 2964 |
| Arg Ala Gly Ala Gly Pro Ser Ser Arg Gly Arg Pro Gly Gly Pro Trp | | | |
| 915 | 920 | 925 | |
| gg ^g ga ^g ag ^c cc ^g tc ^c ag ^t gg ^c cc ^c tc ^c ag ^c ct ^t ga ^g ag ^c ag ^t ga ^g ga ^t | | | 3012 |
| Gly Glu Ser Pro Ser Ser Gly Pro Ser Ser Pro Glu Ser Ser Glu Asp | | | |
| 930 | 935 | 940 | |
| ga ^g gg ^c cc ^a gg ^c cg ^c ag ^c tc ^c ag ^c cc ^c ct ^c cg ^c ct ^g gt ^g cc ^c tt ^c tc ^c | | | 3060 |
| Glu Gly Pro Gly Arg Ser Ser Pro Leu Arg Leu Val Pro Phe Ser | | | |
| 945 | 950 | 955 | |
| ag ^c cc ^g ag ^g cc ^c cc ^c gg ^a ga ^g cc ^g cc ^g gg ^t gg ^g ga ^g cc ^c ct ^g at ^t ga ^g | | | 3108 |
| Ser Pro Arg Pro Pro Gly Glu Pro Pro Gly Glu Pro Leu Met Glu | | | |
| 960 | 965 | 970 | 975 |
| ga ^c tg ^c ga ^g aa ^g ag ^c ag ^c ga ^c ct ^t tg ^c aa ^c cc ^c ct ^g tc ^a gg ^c gg ^c tt ^c | | | 3156 |
| Asp Cys Glu Lys Ser Ser Asp Thr Cys Asn Pro Leu Ser Gly Ala Phe | | | |
| 980 | 985 | 990 | |
| tc ^a gg ^a gt ^g tc ^c aa ^c att tt ^c ag ^c tt ^c tg ^g gg ^g ga ^c ag ^t cc ^g gg ^c cg ^c | | | 3204 |
| Ser Gly Val Ser Asn Ile Phe Ser Phe Trp Gly Asp Ser Arg Gly Arg | | | |
| 995 | 1000 | 1005 | |
| ca ^g ta ^c ca ^g ga ^g ct ^c ct ^c cg ^a tg ^c cc ^c gg ^c cc ^c ac ^c cc ^c ag ^c ct ^c | | | 3249 |
| Gln Tyr Gln Glu Leu Pro Arg Cys Pro Ala Pro Thr Pro Ser Leu | | | |
| 1010 | 1015 | 1020 | |
| ct ^c aa ^c at ^c cc ^c ct ^c tc ^c ag ^c cc ^g gg ^c cc ^c ag ^c gg ^g gg ^c ga ^c | | | 3294 |
| Leu Asn Ile Pro Leu Ser Ser Pro Gly Arg Arg Pro Arg Gly Asp | | | |
| 1025 | 1030 | 1035 | |
| gt ^g ga ^g ag ^c ag ^g ct ^g ga ^t gg ^c ct ^c ca ^g cg ^c ca ^g ct ^c aa ^c ag ^t ct ^g | | | 3339 |
| Val Glu Ser Arg Leu Asp Ala Leu Gln Arg Gln Leu Asn Arg Leu | | | |
| 1040 | 1045 | 1050 | |
| ga ^g cc ^c cg ^g ct ^g ag ^t ga ^c ga ^c at ^g gg ^c ac ^t gt ^c ct ^g ca ^g ct ^g ct ^a | | | 3384 |
| Glu Thr Arg Leu Ser Ala Asp Met Ala Thr Val Leu Gln Leu Leu | | | |
| 1055 | 1060 | 1065 | |
| ca ^g ag ^g ca ^g at ^t ac ^g ct ^g gt ^c cc ^g cc ^c gg ^c tac ag ^t gt ^c gg ^t ac ^c | | | 3429 |
| Gln Arg Gln Met Thr Leu Val Pro Pro Ala Tyr Ser Ala Val Thr | | | |
| 1070 | 1075 | 1080 | |
| ac ^c cc ^g gg ^g ct ^c gg ^c cc ^c ac ^t tc ^c aca tcc cc ^g ct ^g tt ^g cc ^c gt ^c | | | 3474 |
| Thr Pro Gly Pro Gly Pro Thr Ser Thr Ser Pro Leu Leu Pro Val | | | |
| 1085 | 1090 | 1095 | |

Sequence Listing to File.TXT

| | |
|--|------|
| agc ccc ctc ccc acc ctc acc ttg gac tcg ctt tct cag gtt tcc Ser Pro Leu Pro Thr Leu Thr Leu Asp Ser Leu Ser Gln Val Ser | 3519 |
| 1100 1105 1110 | |
| cag ttc atg gcg tgt gag gag ctg ccc ccg ggg gcc cca gag ctt Gln Phe Met Ala Cys Glu Glu Leu Pro Pro Gly Ala Pro Glu Leu | 3564 |
| 1115 1120 1125 | |
| ccc caa gaa ggc ccc aca cga cgc ctc tcc cta ccg ggc cag ctg Pro Gln Glu Gly Pro Thr Arg Arg Leu Ser Leu Pro Gly Gln Leu | 3609 |
| 1130 1135 1140 | |
| ggg gcc ctc acc tcc cag ccc ctg cac aga cac ggc tcg gac ccg Gly Ala Leu Thr Ser Gln Pro Leu His Arg His Gly Ser Asp Pro | 3654 |
| 1145 1150 1155 | |
| ggc agt tagtggggct gcccagtgtg gacacgtggc tcacccaggg atcaaggcgc Gly Ser | 3710 |
| tgctggccg ctccccttgg aggccctgct caggaggccc tgaccgtgga aggggagagg | 3770 |
| aactcgaaag cacagctcct ccccccagccc ttgggaccat cttctcctgc agtccccctgg | 3830 |
| gccccagtga gaggggcagg ggcagggccg gcagtaggtg gggcctgtgg tccccccact | 3890 |
| gccctgaggg cattagctgg tctaactgcc cgaggcacc cggccctggg ccttaggcac | 3950 |
| ctcaaggact tttctgctat ttactgctct tattgttaag gataataatt aaggatcata | 4010 |
| tgaataatta atgaagatgc tgatgactat gaataataaa taattatcct gaggagaaaa | 4070 |

<210> 2
<211> 1159
<212> PRT
<213> Homo Sapiens

<400> 2

Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu Asp
1 5 10 15

Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Ile Ile Ala
20 25 30

Asn Ala Arg Val Glu Asn Cys Ala Val Ile Tyr Cys Asn Asp Gly Phe
35 40 45

Cys Glu Leu Cys Gly Tyr Ser Arg Ala Glu Val Met Gln Arg Pro Cys
50 55 60

Thr Cys Asp Phe Leu His Gly Pro Arg Thr Gln Arg Arg Ala Ala Ala
65 70 75 80

Gln Ile Ala Gln Ala Leu Leu Gly Ala Glu Glu Arg Lys Val Glu Ile
85 90 95

Sequence Listing to File.TXT

Ala Phe Tyr Arg Lys Asp Gly Ser Cys Phe Leu Cys Leu Val Asp Val
100 105 110

Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu Asn
115 120 125

Phe Glu Val Val Met Glu Lys Asp Met Val Gly Ser Pro Ala His Asp
130 135 140

Thr Asn His Arg Gly Pro Pro Thr Ser Trp Leu Ala Pro Gly Arg Ala
145 150 155 160

Lys Thr Phe Arg Leu Lys Leu Pro Ala Leu Leu Ala Leu Thr Ala Arg
165 170 175

Glu Ser Ser Val Arg Ser Gly Gly Ala Gly Gly Ala Gly Ala Pro Gly
180 185 190

Ala Val Val Val Asp Val Asp Leu Thr Pro Ala Ala Pro Ser Ser Glu
195 200 205

Ser Leu Ala Leu Asp Glu Val Thr Ala Met Asp Asn His Val Ala Gly
210 215 220

Leu Gly Pro Ala Glu Glu Arg Arg Ala Leu Val Gly Pro Gly Ser Pro
225 230 235 240

Pro Arg Ser Ala Pro Gly Gln Leu Pro Ser Pro Arg Ala His Ser Leu
245 250 255

Asn Pro Asp Ala Ser Gly Ser Ser Cys Ser Leu Ala Arg Thr Arg Ser
260 265 270

Arg Glu Ser Cys Ala Ser Val Arg Arg Ala Ser Ser Ala Asp Asp Ile
275 280 285

Glu Ala Met Arg Ala Gly Val Leu Pro Pro Pro Pro Arg His Ala Ser
290 295 300

Thr Gly Ala Met His Pro Leu Arg Ser Gly Leu Leu Asn Ser Thr Ser
305 310 315 320

Asp Ser Asp Leu Val Arg Tyr Arg Thr Ile Ser Lys Ile Pro Gln Ile
325 330 335

Thr Leu Asn Phe Val Asp Leu Lys Gly Asp Pro Phe Leu Ala Ser Pro
340 345 350

Sequence Listing to File.TXT

Thr Ser Asp Arg Glu Ile Ile Ala Pro Lys Ile Lys Glu Arg Thr His
355 360 365

Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala Asp Val
370 375 380

Leu Pro Glu Tyr Lys Leu Gln Ala Pro Arg Ile His Arg Trp Thr Ile
385 390 395 400

Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu
405 410 415

Leu Val Ile Tyr Thr Ala Val Phe Thr Pro Tyr Ser Ala Ala Phe Leu
420 425 430

Leu Lys Glu Thr Glu Glu Gly Pro Pro Ala Thr Glu Cys Gly Tyr Ala
435 440 445

Cys Gln Pro Leu Ala Val Val Asp Leu Ile Val Asp Ile Met Phe Ile
450 455 460

Val Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Ala Asn Glu
465 470 475 480

Glu Val Val Ser His Pro Gly Arg Ile Ala Val His Tyr Phe Lys Gly
485 490 495

Trp Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu Ile
500 505 510

Phe Gly Ser Gly Ser Glu Glu Leu Ile Gly Leu Leu Lys Thr Ala Arg
515 520 525

Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser Glu
530 535 540

Tyr Gly Ala Ala Val Leu Phe Leu Leu Met Cys Thr Phe Ala Leu Ile
545 550 555 560

Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Met Glu Gln
565 570 575

Pro His Met Asp Ser Arg Ile Gly Trp Leu His Asn Leu Gly Asp Gln
580 585 590

Ile Gly Lys Pro Tyr Asn Ser Ser Gly Leu Gly Gly Pro Ser Ile Lys
Page 8

Sequence Listing to File.TXT
595 600 605

Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser
610 615 620

Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Ile Phe
625 630 635 640

Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe
645 650 655

Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg
660 665 670

Tyr His Thr Gln Met Leu Arg Val Arg Glu Phe Ile Arg Phe His Gln
675 680 685

Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His Ala
690 695 700

Trp Ser Tyr Thr Asn Gly Ile Asp Met Asn Ala Val Leu Lys Gly Phe
705 710 715 720

Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn Arg Ser Leu
725 730 735

Leu Gln His Cys Lys Pro Phe Arg Gly Ala Thr Lys Gly Cys Leu Arg
740 745 750

Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Pro Pro Gly Asp Thr
755 760 765

Leu Val His Ala Gly Asp Leu Leu Thr Ala Leu Tyr Phe Ile Ser Arg
770 775 780

Gly Ser Ile Glu Ile Leu Arg Gly Asp Val Val Val Ala Ile Leu Gly
785 790 795 800

Lys Asn Asp Ile Phe Gly Glu Pro Leu Asn Leu Tyr Ala Arg Pro Gly
805 810 815

Lys Ser Asn Gly Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His Lys
820 825 830

Ile His Arg Asp Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Glu Phe
835 840 845

Sequence Listing to File.TXT

Ser Asp His Phe Trp Ser Ser Leu Glu Ile Thr Phe Asn Leu Arg Asp
850 855 860

Thr Asn Met Ile Pro Gly Ser Pro Gly Ser Thr Glu Leu Glu Gly Gly
865 870 875 880

Phe Ser Arg Gln Arg Lys Arg Lys Leu Ser Phe Arg Arg Arg Thr Asp
885 890 895

Lys Asp Thr Glu Gln Pro Gly Glu Val Ser Ala Leu Gly Pro Gly Arg
900 905 910

Ala Gly Ala Gly Pro Ser Ser Arg Gly Arg Pro Gly Gly Pro Trp Gly
915 920 925

Glu Ser Pro Ser Ser Gly Pro Ser Ser Pro Glu Ser Ser Glu Asp Glu
930 935 940

Gly Pro Gly Arg Ser Ser Ser Pro Leu Arg Leu Val Pro Phe Ser Ser
945 950 955 960

Pro Arg Pro Pro Gly Glu Pro Pro Gly Gly Glu Pro Leu Met Glu Asp
965 970 975

Cys Glu Lys Ser Ser Asp Thr Cys Asn Pro Leu Ser Gly Ala Phe Ser
980 985 990

Gly Val Ser Asn Ile Phe Ser Phe Trp Gly Asp Ser Arg Gly Arg Gln
995 1000 1005

Tyr Gln Glu Leu Pro Arg Cys Pro Ala Pro Thr Pro Ser Leu Leu
1010 1015 1020

Asn Ile Pro Leu Ser Ser Pro Gly Arg Arg Pro Arg Gly Asp Val
1025 1030 1035

Glu Ser Arg Leu Asp Ala Leu Gln Arg Gln Leu Asn Arg Leu Glu
1040 1045 1050

Thr Arg Leu Ser Ala Asp Met Ala Thr Val Leu Gln Leu Leu Gln
1055 1060 1065

Arg Gln Met Thr Leu Val Pro Pro Ala Tyr Ser Ala Val Thr Thr
1070 1075 1080

Pro Gly Pro Gly Pro Thr Ser Thr Ser Pro Leu Leu Pro Val Ser
1085 1090 1095

Sequence Listing to File.TXT

Pro Leu Pro Thr Leu Thr Leu Asp Ser Leu Ser Gln Val Ser Gln
1100 1105 1110

Phe Met Ala Cys Glu Glu Leu Pro Pro Gly Ala Pro Glu Leu Pro
1115 1120 1125

Gln Glu Gly Pro Thr Arg Arg Leu Ser Leu Pro Gly Gln Leu Gly
1130 1135 1140

Ala Leu Thr Ser Gln Pro Leu His Arg His Gly Ser Asp Pro Gly
1145 1150 1155

Ser

<210> 3
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 3
aatttgttacc atgggctca tagatgccgggtc c 31

<210> 4
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 4
gcttttacttc aggcagcacg t 21

<210> 5
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 5
ccaccaggta ccgtgagatc a 21

<210> 6
<211> 21
<212> DNA
<213> Artificial

<220>

Sequence Listing to File.TXT

<223> primer

<400> 6

ttgcagtgct gcagcagtga g

21

<210> 7

<211> 21

<212> DNA

<213> Artificial

<220>

<223> primer

<400> 7

atgctagcat cttcggcaac g

21

<210> 8

<211> 31

<212> DNA

<213> Artificial

<220>

<223> primer

<400> 8

aattaagctt tttcgagttc ctctccctt c

31

<210> 9

<211> 52

<212> DNA

<213> Artificial

<220>

<223> oligo DNA

<400> 9

gatccccccgg gctgcaggaa ttgcataatcg ttaacgtcga cctcgagggt ac

52

<210> 10

<211> 44

<212> DNA

<213> Artificial

<220>

<223> oligo DNA

<400> 10

cctcgagggtc gacgttaacg atatcgaatt cctgcagccc gggg

44

<210> 11

<211> 32

<212> DNA

<213> Artificial

<220>

<223> primer

<400> 11

Sequence Listing to File.TXT

gtcgtcatcg atacaaatgg cagtattcat cc

32

<210> 12
<211> 34
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 12
gtcgtcaagc ttccaaactg gatctctgct gtcc

34

<210> 13
<211> 5020
<212> DNA
<213> Retroviral provirus

| | |
|---|------|
| <400> 13 | |
| ctgcagcctg aatatggccc aaacaggata tctgtggtaa gcagttcctg ccccggtca | 60 |
| gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt | 120 |
| tcctgccccg gctcagggcc aagaacagat ggtccccaga tgccgtccag ccctcagcag | 180 |
| tttcttagaga accatcagat gtttccaggg tgccccaaagg acctgaaatg accctgtgcc | 240 |
| ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgccgcgttc tgctccccga | 300 |
| gctcaataaaa agagcccaca acccctcaact cggggcgcca gtcctccgat tgactgagtc | 360 |
| gcccgggtac ccgtgtatcc aataaaaccct cttgcagttg catccgactt gtggtctcgc | 420 |
| tgttccttgg gagggctcgtc tctgagtgtat tgactacccg tcagcggggg tctttcattt | 480 |
| gggggctcgt ccgggatcgg gagacccctg cccagggacc accgaccac caccgggagg | 540 |
| taagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgattttta | 600 |
| tgcgcctgcg tcggacttag ttagctaact agctctgtat ctggcggacc cgtggtgaa | 660 |
| ctgacgagtt ctgaacacccc ggccgcaacc ctgggagacg tcccaggac tttggggcc | 720 |
| gtttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg | 780 |
| tggttcttgtt aggagacgag aacctaataac agttcccgcc tccgtctgaa tttttgtttt | 840 |
| cggtttgaa ccgaagccgc gcgtcttgc tgctgcagca tcgttctgtg ttgtctctgt | 900 |
| ctgactgtgt ttctgtatTT gtctgaaaat tagggccaga ctgttaccac tcccttaagt | 960 |
| ttgacccctag atcactggaa agatgtcgag cggctcgctc acaaccagtc ggttagatgtc | 1020 |
| aagaagagac gttgggttac cttctgctct gcagaatggc caacctttaa cgtcggatgg | 1080 |
| ccgcgagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa ggtctttca | 1140 |
| cctggccgc atggacacccc agaccaggtc ccctacatcg tgacctggaa agccttggct | 1200 |
| tttgacccccc ctccctgggt caagccctt gtacacccta agcctccgccc tcctttctt | 1260 |

Sequence Listing to File.TXT

| | | | |
|------------------------|------------------------|-------------------------|------|
| ccatccgcgc cgtctctccc | ccttgaacct cctcttcga | ccccgcctca atcctccctt | 1320 |
| tatccagccc tcactccttc | tctaggcgcc ggccggatcc | cagtgtggtg gtacgttagga | 1380 |
| attcgccagc acagtggtcg | acctgtggaa tgtgtgtcag | ttagggtgtg gaaagtcccc | 1440 |
| aggctccccca gcaggcagaa | gtatgcaaag catgcatctc | aattagtcag caaccaggtg | 1500 |
| tggaaagtcc ccaggctccc | cagcaggcag aagtatgcaa | agcatgcata tcaattagtc | 1560 |
| agcaaccata gtcccgc(cc | taactccgccc catcccgc | ctaactccgc ccagttccgc | 1620 |
| ccattctccg cccatggct | gactaatttt ttttatttat | gcagaggccg aggccgcctc | 1680 |
| ggcctctgag ctattccaga | agtagtgagg aggcttttt | ggaggcctag gctttgcaa | 1740 |
| acgctgcttgg | aggctgaagg tgcgttgctg | gcgttttcc ataggctccg | 1800 |
| gagcatcaca aaaatcgacg | ctcaagtca aggtggcgaa | acccgacagg actataaaga | 1860 |
| taccaggcgt ttccccctgg | aagctccctc gtgcgctctc | ctgttccgac cctgccgctt | 1920 |
| accggataacc tgtccgcctt | tctccctcg ggaagcgtgg | cgctttctca tagtcacgc | 1980 |
| tgttaggtatc tcagttcggt | gtaggtcggt cgctccaagc | tggctgtgt gcacgaaccc | 2040 |
| cccggtcagc ccgaccgctg | cgccttatcc ggttaactatc | gtcttgagtc caacccggta | 2100 |
| agacacgact tatcgccact | ggcagcagcc actggtaaca | ggattagcag agcgaggtat | 2160 |
| gtaggcggtg ctacagagtt | cttgaagtgg tggcctaact | acggctacac tagaaggaca | 2220 |
| gtatttggta tctgcgtct | gctgaagcca gttaccttcg | aaaaaagagt tggtagctct | 2280 |
| tgatccggca aacaaaccac | cgctggtagc ggtggtttt | ttgtttgcaa gcagcagatt | 2340 |
| acgatcgata aaataaaaga | ttttattn tctccagaaa | aaggggggaa tgaaagaccc | 2400 |
| cacctgtagg tttggcaagc | tagcttaagt aacgccattt | tgcaaggcat ggaaaaatac | 2460 |
| ataactgaga atagagaagt | tcagatcaag gtcaggaaca | gatggAACAG ctgaatatgg | 2520 |
| gccaacagg atatctgtgg | taagcagttc ctgccccggc | tcagggccaa gaacagatgg | 2580 |
| aacagctgaa tatgggccaa | acaggatatc tgtggtaagc | agttcctgcc ccggctcagg | 2640 |
| gccaagaaca gatggtcccc | agatgcggc cagccctcag | cagtttctag agaaccatca | 2700 |
| gatgtttcca gggtgccca | aggacctgaa atgaccctgt | gccttatttg aactaaccaa | 2760 |
| tcagttcgct tctcgcttct | gttcgcgcgc ttctgctccc | cgagctcaat aaaagagccc | 2820 |
| acaaccctc actcggggcg | ccagtcctcc gattgactga | gtcgccccggg taccctgtta | 2880 |
| tccaataaac cctcttgca | ttgcattccga cttgtggct | cgctgttccct tgggagggtc | 2940 |
| tcctctgagt gattgactac | ccgtcagcgg gggtcttca | catgcagcat gtatcaaataat | 3000 |
| taatttggtt tttttctta | agtatttaca ttaaatggcc | atagttgcat taatgaatcg | 3060 |
| gccaacgcgc ggggagagggc | ggtttgcgtt ttgggcgtc | ttccgcttcc tcgctcactg | 3120 |
| actcgctgcg ctcggtcgtt | cggtcgccg gagcggatcc | agctcactca aaggcggtaa | 3180 |

Sequence Listing to File.TXT

| | |
|--|------|
| tacggttatc cacagaatca ggggataacg cagggaaagaa catgtgagca aaaggccagc | 3240 |
| aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggtt tttccatagg ctccgc | 3300 |
| ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat | 3360 |
| aaagatacca ggcgtttccc cctggaagct ccctcgtcg ctctcctgtt ccgaccctgc | 3420 |
| cgcttaccgg atacctgtcc gccttctcc cttcggaaag cgtggcgctt tctcatagct | 3480 |
| cacgctgttag gtatctcagt tcggtgttagg tcgttcgctc caagctggc tgtgtgcacg | 3540 |
| aaccccccgt ttagccccac cgctgcgc tatccggtaa ctatcgctt gagtccaacc | 3600 |
| cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga | 3660 |
| ggtatgttagg cggtgctaca gagttcttga agtggtgcc taactacggc tacactagaa | 3720 |
| ggacagtatt tggtatctgc gctctgctga agccagttac cttcgaaaaa agagttggta | 3780 |
| gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt tgcaagcagc | 3840 |
| agattacgcg cagaaaaaaaaa ggatctcaag aagatccccc gatctttct acggggctcg | 3900 |
| acgctcagtgcg gaacgaaaac tcacgttaag ggatttttgtt catgagatta tcaaaaagga | 3960 |
| tcttcaccta gatcctttta aattaaaaat gaagtttgcg gccgcaaatc aatctaaagt | 4020 |
| atatatgagt aaacttggtc tgacagttac caatgctta tcagtggggc acctatctca | 4080 |
| gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcggtta gataactacg | 4140 |
| atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca | 4200 |
| ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtgg | 4260 |
| cctgcaactt tatccgcctc catccagtct attaattttt gcccggaaagc tagagtaagt | 4320 |
| agttcgccag ttaatagtt gcgcaacgtt gttgccatttgcg ctacaggcat cgtgggtca | 4380 |
| cgctcgctgt ttggatggc ttcatcagc tccgggtttccc aacgatcaag gcgagttaca | 4440 |
| tgtatccccca tgggttgcaaa aaaagcggtt agtccttcg gtcctccgat cggtgtcaga | 4500 |
| agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact | 4560 |
| gtcatgccat ccgtaagatg ctttctgtg actgggtgagt actcaaccaa gtcattctga | 4620 |
| gaatagtgtt tgcggcgacc gagttgtct tggccggcgtaa caacacggtaa taataccgcg | 4680 |
| ccacatagca gaactttaaa agtgcgtatc attggaaaac gttttcgaaa gcgaaaaactc | 4740 |
| tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgatc acccaactga | 4800 |
| tcttcagcat cttttactttt caccacgtt tctgggtgag caaaaacagg aaggcaaaat | 4860 |
| gccgcaaaaa aggaaataag ggcgacacgg aaatgttcaa tactcataact cttccttttt | 4920 |
| caatattattttaa gaagcattta tcagggttat tgcgtatgcgatc acatatttgcgat | 4980 |
| attttagaaaaa ataaacaaat aggggttccg cgcacatttc | 5020 |